Uncertainty Factors in Ecotoxicological Risk Management

Peter Craig
University of Durham

March 9, 2006

Abstract

Uncertainty factors (also known as extrapolation or safety factors) are widely used in lower tier ecotoxicological risk management in order to allow for sources of variability and uncertainty for which there is only limited information. From the point of view of probabilistic risk assessment, the rationale underlying both the use of uncertainty factors and the specific values chosen is often unclear. On the other hand, uncertainty factors are a convenient approach to lower tier risk assessment.

This paper considers the problem of how to improve these lower tier risk assessments without moving completely away from the use of uncertainty factors towards complex probabilistic modelling which often requires calculations using Monte Carlo. It is shown how relatively simple statistical models lead to manageable procedures for computing a single uncertainty factor for multiple sources of uncertainty including inter-species variation, elaborating on ideas and methodology presented in a recent European Food Safety Authority scientific opinion in relation to aquatic risk assessment for pesticides.

1 Introduction

This paper was stimulated by the author's involvement as a so-called *ad hoc expert* providing research and advice to the PPR (plant health, plant protection products and their residues) panel of EFSA (European Food Safety Authority) in relation to an EFSA scientific opinion ([1]) on uncertainty factors for use in aquatic ecotoxicology. The views expressed here should not be taken to represent EFSA, the members of the panel or the other ad hoc experts involved with the opinion.

In many areas of ecological and environmental risk management for hazardous substances, the application of one or more uncertainty factors to a statistic derived from relevant measurements is the standard method for determining a regulatory threshold in lower-tier risk management. Relevant government agency documents include [2], [3], [4] and [5].

The context in [1] is risk assessment for effects of pesticides on aquatic ecosystems. The risk management problem is the licensing of pesticides for agricultural use. The standard regulatory approach is to require a comparison of predicted exposure concentration (under planned use of pesticide) to some suitable measure of acceptable exposure concentration (known hereafter as AEC) derived from tests of sensitivity of individual species. In general, the AEC is supposed the be a level at which sufficiently adverse consequences are unlikely to occur; in some contexts the desired level is the so-called predicted no-effect concentration (PNEC). If predicted exposure is less than the AEC, the pesticide is passed for use (in reality there are also assessments for other risks which must be passed). If predicted exposure is greater than the AEC then the pesticide can only be licensed by undertaking a higher-tier risk assessment which is generally a much more detailed and expensive assessment of the nature and magnitude of risk involved. As usual, the idea is that the lower tier procedure is reasonably cautious and that a

pesticide which passes at that stage is very unlikely to cause problems. However, it's not at all clear how this is calibrated given the lack of transparency and apparent lack of rationality (from a probabilistic perspective) in the approach which is used to derive the AEC.

In this paper, the focus is on simple procedures for determining a suitable AEC for a new substance (such as a pesticide). The goal is not to develop a fully believable probabilistically coherent approach. In principle, risk should be assessed in terms of the final risk management decision. However, the science of toxicity is quite distinct from that underlying exposure assessment and the experts on one are not usually experts on the other. Moreover, toxicity by itself presents plenty of obstacles to overcome. Simple procedures are needed for routine risk management especially in a world where risk managers are not themselves highly quantitatively literate. It is stated in [6] that, despite the existence for the previous decade of a legal requirement for risk assessment for industrial and agricultural chemicals, the process had been carried out for fewer than 10 of the 100,000 or so chemicals registered by industry. The EFSA opinion on aquatic ecotoxicology ([1]) was constrained by the need to provide computationally straightforward procedures requiring very little, if any, mathematical sophistication from the user. The need for simple procedures means that the statistical models used must themselves be simple but it is nevertheless important that they capture the key features of the problem.

In current practice for the aquatic context, the procedure for computing the AEC is to start by testing a number of species for sensitivity to the pesticide, the results being expressed in terms of exposure concentrations. Then the value obtained for the most sensitive species is divided by one or more uncertainty factors (also known as extrapolation, safety or assessment factors) to obtain the AEC. A useful critical discussion of the use of uncertainty factors in ecological risk assessment may found in [6].

In the literature on quantitative methods for ecological risk assessment (for example, see [7] and [8]), the usual underlying measure of performance is the proportion of species suffering adverse effects if exposed to a concentration below or equal to the AEC. Clearly, it would be better to use a measure of ecosystem response. However, that would be more difficult both to define and to measure; moreover, there may not be any clear ecosystem or group of ecosystems to be considered at the lower tiers of risk assessment. This does not preclude the use of an uncertainty factor to adjust for differences between consequences for an ecosystem and for individual species but such a factor is not included in the proposals for risk calculations given later and would have to be applied separately.

2 Uncertainty factors

In many areas of application, individual uncertainty factors are not clearly specified but the reasoning presented in the literature indicates that overall uncertainty factors are usually obtained by multiplying individual factors related to specific sources of uncertainty. A description of factors recommended for use in the European Union and by the US Office of Pollution Prevention and Toxics may be found in [6].

The glossary ([9]) for the US Environmental Protection Agency's Integrated Risk Information System refers to separate uncertainty factors for intra-species variability, interspecies uncertainty, extrapolation from sub-chronic to chronic exposure, extrapolating from low-effect data to no-effect and uncertainty associated with an incomplete database. Moreover the US Food Quality Protection Act ([10]) requires the use of an additional uncertainty factor to account for possible differences in sensitivity between adults and children.

A detailed discussion of the history of uncertainty factors used in aquatic ecotoxicology is given in [1] and it is clear that the overall uncertainty factor is the product of individual factors. In aquatic ecotoxicology, uncertainty factors are applied to allow for

the difference between measurements under laboratory conditions of some suitable exposure concentration for a toxicity endpoint such as LC_{50}[1] over some suitable time period and the exposure concentration which will result in significant adverse consequences in the ecosystem. Issues being covered by uncertainty factors include differences between acute and chronic sensitivity, between laboratory and field environments and between individual species and ecosystems.

On the one hand, uncertainty factors are a simple way to account for multiple sources of uncertainty in the relationship between laboratory data and adverse consequences in the system being managed. On the other hand, the choice of value for an individual uncertainty factor is rarely given a detailed justification and the risk resulting from applying multiple factors is poorly quantified.

A relatively simple approach to the problem is to model the sources of uncertainty probabilistically as has been attempted in in [11] and [12]; the latter also includes a number of references to earlier related work. However, the authors of [12] clearly feel the need to derive the distributions used purely from empirical evidence rather than expert judgment and [11] proposes an inappropriate method for combining factors. Both [11] and [12] assume that the sources may be treated independently, an assumption which is convenient but not really necessary.

To simplify matters, we will focus on the two extreme ends of the process of extrapolation for a particular substance. Let l denote the laboratory measurement for a species of the exposure concentration corresponding to some suitable (probably acute) toxicity endpoint and f denote the unknown/unknowable exposure concentration for which there would be no adverse consequences in the field. The difference between f and l originates from a number of sources which are accounted for by uncertainty factors. The way in which the uncertainty factors are applied and the literature on their use strongly suggests a hypothesis of additive components of uncertainty on logarithmic[2] scale, i.e. in relation to $\log l$ and $\log f$. First, since the AEC is obtained by dividing a value of l by uncertainty factors, $\log \mathrm{AEC}$ is obtained by subtracting uncertainty shifts from $\log l$. Secondly, the magnitude of l does not affect the factors applied. Thirdly, data are nearly always statistically modelled and analysed using $\log l$.

In [13], an analysis is given of the relationship between exposure concentrations corresponding to particular acute and chronic toxicity endpoints for a substantial number of substances and species for which both forms of endpoint had been measured in the same experiment. The substances are categorised into four chemical groups and a separate linear regression is carried out for each group to model the dependence of $\log f$ on $\log l$. Interestingly, the hypothesis that the slope is 1 in each regression is not explored. The data were reexamined for this paper and for three of the groups the data are consistent with a slope of 1; for the fourth group it can be argued that the slope is statistically significantly different from 1 but there would be little loss of predictive accuracy in using a slope of 1. Figure 1 shows the data for each group with a line overlaid having slope 1. Thus a simple model is that the difference between $\log f$ and $\log l$ is distributed independently of the laboratory measurement. The distribution of the difference was found to be reasonably normal although there is some suggestion of a long tail at lower values which merits further investigation as it is potentially important from risk management perspective. The mean and standard deviation of the difference were found to vary significantly from one group of substances to another.

There is not a large amount of readily available data to be used for statistical modelling of other sources of difference between $\log l$ and $\log f$. This does not preclude the use of expert judgment/opinion. After all, current values for uncertainty factors arise in

[1] LC_{50} is the concentration at which 50% of individuals die in the specified time period

[2] In the toxicity and ecological risk assessment communities, logarithms are always base 10

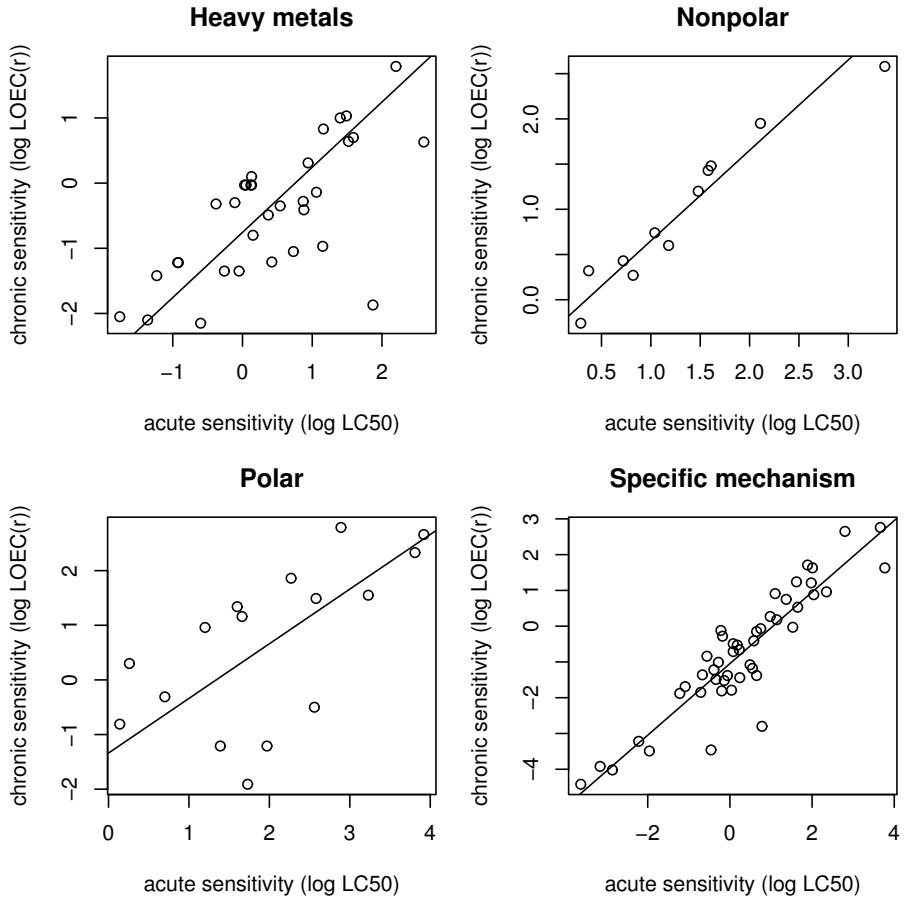

Figure 1: Data from [13] on relationship between chronic and acute sensitivity for four classes of chemical (heavy metals, polar, non-polar, specific mechanism). Slope of overlaid line in each panel is 1.

that way. However, current uncertainty factors do not distinguish mean and standard deviation which play different roles; the mean acts as a bias term to change the typical exposure concentration whereas the standard deviation encodes additional variability. A single uncertainty factor confounds these two effects. Of course, experts have to specify just one number for an uncertainty factor as opposed to two for a mean and standard deviation. On the other hand, it is difficult to understand how a careful choice of uncertainty factor could be made without first considering the two issues of typical change in exposure concentration and magnitude of variability in the change.

Suppose then that the difference between $\log l$ and $\log f$ is the sum of a number of components each of which has a normal distribution and that a mean and standard deviation is known for each component. Then of course the distribution of the difference is itself normal with mean μ_d being the sum of the component means and standard deviation σ_d being the square root of the sum of the squares of the individual standard deviations. In the event that data or expert judgment suggest the components are not independent it is simple to incorporate correlations between the components.

Clearly there is no reason why each source of variability should have a normal distribution or why they should all be purely additive (on log-scale). However, the goal here is to develop simple procedures for risk management and the assumptions being made here do lead to tractable calculations. If necessary, the means and standard deviations may need to be specified conservatively to allow for departures from normality

or uncertainties about their values if only limited data are available.

3 Species sensitivity distributions and risk quantification

In recent years, the concept of species sensitivity distribution has acquired a lot of support in ecological risk assessment. For a comprehensive overview, see [14]. The idea makes many possible many quantitative risk calculations although it has been criticised, for example in [6].

The basic idea of the species sensitivity distribution (SSD) is that for a given substance and chosen measure of sensitivity[3] (such as $\log l$), there is a distribution of measured sensitivities between species. No distinction is usually made between actual and measured sensitivity, perhaps because of difficulties in quantifying measurement error, especially in databases of results from several decades of experiments. In practice, one tends to assume that the geometric mean (mean on log-scale) of all the relevant data is the actual sensitivity. The role of the SSD is to act as a surrogate for all the individual species sensitivities for the species one would find in the ecosystem(s) under consideration. In reality, each ecosystem is unique and the number of species involved is finite though possibly large; theoretically, it is convenient to assume a continuous distribution of sensitivity which applies to all ecosystems being considered. Because there are different kinds of sensitivity, different ways to measure them and possibly different laboratories to consider, there are in fact many SSDs for any given substance and population of species. For the purposes of this paper we will confine attention to just two: the SSD for $\log l$ and the SSD for $\log f$.

However, there is a second hidden assumption which permeates the SSD literature. Implicitly it is assumed that, for a new substance, the sensitivities of the species represented by the SSD are a priori exchangeable. At the same time, it is fairly widely accepted that some species are more sensitive than others; indeed, the legislative requirement for the use of rainbow trout in aquatic ecotoxicology is rooted partly in the relative ease of measurement of sensitivity for this species and partly in the knowledge that it tends to be more sensitive than average (see [1]).

Once one accepts that a single species is more sensitive than others, one should accept the possibility that there are fundamental differences in sensitivity for many (if not all) species. This would appear to totally undermine the SSD concept. However, it depends on the view one takes of the role of the SSD in risk assessment and management. If the role is to act as (foundation for) the predictive distribution for sensitivity of un-measured species on the basis of measurements for others, it may be possible to salvage the concept. An important realisation is that, once we have per-species distributions, there is an implicit assumption of some population of substances, not just some population of relevant species. In fact this same assumption underlies much of the SSD literature but is not emphasised.

In [1] it is assumed throughout that the SSD is normal but in certain situations applies to all except one species which is given a different (implicitly subjective/Bayesian predictive) normal distribution with mean and standard deviation obtained from (but differing from) those of the SSD. The problem is how to extend that idea to allow for the possibility that there are many (possibly all) species which are typically more or less sensitive than the average. The difficulty is that it is no longer easy in practice to define the distribution of individual species by reference to the common distribution of the remainder of species. Indeed it is easy to end up with a probabilistically incoherent model. It is certainly not clear how to do so while maintaining tractability of the resulting

[3] Sensitivity is really a misnomer for $\log l$ as higher values of the latter correspond to less sensitive species. However, this is the standard usage.

risk calculations.

The pragmatic approach offered here (and to a limited extent in [1]) is to retain exchangeability of sensitivity for all species other than those whose sensitivity is measured for the risk assessment; thus the SSD describes those species. We then measure the performance of the choice of AEC in terms of the SSD. The presumption is that whatever procedure we adopt for determining the AEC, the measured species will themselves be adequately protected.

Adapting the approach in [1], the adverse consequences of a particular choice of AEC are measured, using the SSD for $\log f$, as the proportion of species for which the field sensitivity $\log f$ of the species lies below $t = \log AEC$, in other words it is the cumulative distribution function (cdf) of the field SSD evaluated at t. For the sake of subsequent mathematical simplicity we assume that the SSD for $\log l$ is normal with mean μ and standard deviation σ and so the SSD for $\log f$ is normal mean $\mu_f = \mu + \mu_d$ and variance $\sigma_f^2 = \sigma^2 + \sigma_d^2$. Thus if t, μ and σ are known, the adverse consequences are measured by $\Phi\big((t - \mu_f)/\sigma_f\big)$ where $\Phi(.)$ denotes the cdf of the standard normal distribution.

To cope with uncertainty about μ and σ (and t when considering the frequentist viewpoint), the expected adverse consequences (risk in the statistical sense) are measured by $\mathrm{E}[\,\Phi\big((t - \mu_f)/\sigma_f\big)\,]$ where the expectation is computed using respectively the sampling distribution of t and the posterior distribution of μ and σ when considering the frequentist and Bayesian viewpoints. It can be argued that this measure is too crude as it only measures the average/expected proportion of species "at risk" and not variability/uncertainty in that proportion; however, the Markov inequality guarantees that a high proportion at risk is unlikely if the expectation is made small and hopefully that will hold in application.

An alternative way of understanding this measure is that it is the probability that randomly chosen value y from the field SSD lies below t. From the frequentist perspective, the probability is evaluated taking both t and y as random whereas a Bayesian evaluates the probability using the posterior predictive distribution for y treating t as constant.

4 Statistical modelling

The data to be used will be values of l measured for n species. Denote the resulting values of $\log l$ by $x_1, \ldots, x_n$. Denote the corresponding unknown values of $\log f$ by $y_1, \ldots, y_n$ where the relationship between y and x is that $y_i = x_i + d_i$ and $d_1, \ldots, d_n$ are independent and identically distributed with known mean μ_d and standard deviation σ_d.

In order to arrive at simple rules for determining the AEC, we assume that the variation in σ_d from substance to substance exactly follows the variation in σ between substances, i.e. $\sigma_d = \beta\sigma$ for some known value β. This is a very strong assumption which clearly should be validated using data. However, the data in [13] only provide information about d and not about the value of σ. Moreover the database of fish sensitivities used in [1] is confidential and permission has not been obtained to use it here. If this assumption fails, a more sophisticated model of the relationship between σ_d and σ will be needed and even a simple linear relationship between the variances presents considerable extra challenges in terms of deriving simple procedures for determining the AEC.

4.1 Non-exchangeability

In [1], non-exchangeability is incorporated by the assumption that x_i is normal with mean $\mu - k_i\sigma$ and standard deviation $\phi_i\sigma$ where k_i represents the "bias" of species i and ϕ_i allows for different variability for species i. Implicitly, k_i and ϕ_i are defined with reference to some suitable population of substances and are assumed known. The idea

of using $k_i\sigma$ as the adjustment to the mean was to allow the expected position of the species in the SSD to be unaffected by variability in σ from substance to substance. A method was provided by the author in [15] for estimating k_i and ϕ_i for a single biased species, from a suitable database of measured sensitivities for many substance/species combinations, using maximum likelihood estimation based on appropriate t-statistics. The method extends straightforwardly to more than one biased species but will become increasingly and undesirably sensitive to the structure of the database as the number of such species increases.

However, in [1], non-exchangeability was not pursued for methods of determining the AEC for a specified level of adverse consequences but only for those special methods which sought to maintain the current unspecified level of environmental protection without unnecessarily discouraging notifiers from measuring more species than the minimum required by regulations. A slight modification to the non-exchangeability model leads to our desired goal of simple rules for the determining the AEC. Rather than taking the mean of x_i to be $\mu - k_i\sigma$, we take it to be $\mu - k_i$. This does mean that the expected position of the species in the SSD varies from substance to substance with σ; however, the data on for acute toxicity in the RIVM database used in [1] do not strongly suggest that the adjustment should be proportional to σ.

Standard linear model theory (see [16]) then shows that the best estimate of μ is $\hat{\mu} = [\sum_i \phi_i^{-2}(x_i + k_i)]/[\sum_i \phi_i^{-2}]$ and the best estimate of σ^2 is s^2 where $(n-1)s^2 = \sum_i \phi_i^{-2}(x_i + k_i - \hat{\mu})^2$. Moreover, $\hat{\mu}$ is normal with mean μ and variance $\sigma^2/[\sum \phi_i^{-2}]$ and $(n-1)s^2/\sigma^2$ has the chi-squared distribution with $n-1$ degrees of freedom.

4.2 Borrowing strength from other substances

Where toxicity data are available for other relevant substances, it may be possible to derive information about the distribution of σ from substance to substance. Two simple situations are considered in [1].

In the first situation, it is supposed that there is a large database relating to substances which might be considered to be exchangeable with the substance being assessed. The database may be used to estimate the shape parameter α and rate parameter λ for a gamma-distribution modelling variability of σ^{-2} from substance to substance; the estimation procedure given in [15] is based on maximising a suitable marginal likelihood for the parameters. Then we suppose that σ^{-2} for the substance being assessed is drawn from this gamma-distribution (equivalently, consider it to be a prior for σ^{-2} in the Bayesian viewpoint). This leads to an improved estimator of σ^2: $s_A^2 = [2\lambda + (n-1)s^2]/[2\alpha + (n-1)]$ where $\big(2\alpha + (n-1)\big)s_A^2/\sigma^2$ is chi-squared with $2\alpha + (n-1)$ degrees of freedom.

In the second situation, it is supposed that toxicity data $\log l$ are available for a small number N of other substances for each of which there is good reason to believe that the standard deviation will be very similar to that of the substance being assessed. Let n_j be the number of observations for the jth other substance and s_j^2 the estimate of σ^2 from the data for that substance. Then this leads to the usual pooled variance estimate $s_P^2 = [(n-1)s^2 + \sum_j (n_j - 1)s_j^2]/[(n-1) + \sum_j (n_j - 1)]$ having $(n-1) + \sum_j (n_j - 1)$ degrees of freedom.

5 Determining the AEC

In what follows, S denotes s, s_A or s_P and m denotes the associated degrees of freedom so that mS^2/σ^2 is chi-squared with m degrees of freedom.

The obvious way to define the AEC is to put

$$t = \hat{\mu} + \mu_d - kS \tag{1}$$

for some appropriate choice of k to be determined to achieve the required level of statistical risk. In effect, we are computing a single uncertainty factor covering inter-species variation as well as extrapolation from laboratory to field. The uncertainty factor is simply $10^{kS-\mu_d}$ and is being applied to a weighted geometric-type mean of $l_1, \ldots, l_n$ after an adjustment for non-exchangeability.

The frequentist computation is $\mathrm{P}[y < t] = \mathrm{P}[(y-\hat{\mu}-\mu_d)/S < -k]$. However $y-\hat{\mu}-\mu_d$ is normal with mean 0 and variance $(\psi\sigma)^2$ where $\psi^2 = (1 + \beta^2 + 1/[\sum_i \phi_i^{-2}])$ so that $(y - \hat{\mu} - \mu_d)/(\psi S) = T_m$ where T_m has the t-distribution with m degrees of freedom. Hence if p is the required value of $\mathrm{P}[y < t]$, $k = \psi t_{m,p}$ where $t_{m,p}$ is the $100(1-p)$th percentile of the t-distribution having m degrees of freedom.

From the Bayesian perspective, standard Bayesian theory for the linear model (see for example [17]) shows that the posterior predictive distribution for $(y - \hat{\mu} - \mu_d)/(\psi S)$ is the t-distribution with m degrees of freedom, provided we use the prior distribution $p(\tau, \mu) \propto 1/\tau$ when there is no available data from other relevant substances, $p(\tau, \mu) \propto \tau^{\alpha-1} \exp\{-\lambda\tau\}$ in the first situation discussed in section 4.2, and $p(\tau, \mu, \mu_1, \ldots, \mu_N) \propto 1/\tau$ in the second situation. Consequently the Bayesian statistical risk of (1) is exactly the same as the frequentist statistical risk in these circumstances. Thus provided $k = \psi t_{m,p}$, (1) defines the Bayesian decision rule.

Given the model, the Bayesian has no choice about how to define t whereas the frequentist could for example instead choose to base t on some suitable order statistic of the data; for other reasons, this was the approach taken in [1] for one particular situation. The fact that (1) is the Bayesian solution provides strong evidence that this is an efficient choice of rule from the frequentist point of view.

6 Conclusions and future directions

We have seen that it is possible to arrive a straightforward procedure for finding the AEC. In effect, it involves applying a single uncertainty factor to an appropriate summary of laboratory toxicity data. Unlike standard factors currently used, the proposed uncertainty factor explicitly depends on the mean and standard deviation of the difference between laboratory and field sensitivity, on information concerning non-exchangeable species, on the number of species tested, and on an estimate of the standard deviation of the laboratory SSD derived from available toxicity data.

However, there are are several very strong modelling assumptions being made, both distributionally and otherwise, in order to achieve a simple procedure. Where possible, these should be validated using data and otherwise sensitivity of the result should be examined. Moreover, the possibility of weakening the assumptions at the price of small increases in complexity of the procedure should be investigated. Of these assumptions, the most obvious weakness is the assumption that $\sigma_d \propto \sigma$.

Even where quite large databases of toxicity data exist, the combinations of substance and species tested have arisen largely by historical accident and are not necessarily suitable for estimating non-exchangeability parameters or differences between various toxicity measures. There is an interesting experimental design problem to decide how best to expend some further experimental effort to maximise the return from the databases.

It's fairly obvious that the most important aspects of uncertainty are those for which it's most difficult and expensive to obtain quantitative data, namely the sources of differences between field and laboratory sensitivities: μ_d and β. Obtaining information about the shape of the distribution of $\log l$, variability of σ and about non-exchangeability of $\log l$ is much easier but ultimately less informative.

Acknowledgments I would like to thank EFSA, Tom Aldenberg, Andy Hart, Robert Luttik, Willem Roelofs and the members of the PPR panel for introducing me to the area of ecotoxicological risk assessment, providing lots of data, insight and ideas and putting up with the results.

References

[1] European Food Safety Authority panel on Plant Health, Plant Protection Products and their Residues, *Opinion of the PPR Panel on a request from EFSA related to the assessment of the acute and chronic risk to aquatic organisms with regard to the possibility of lowering the uncertainty factor if additional species were tested*, The EFSA Journal, **301**, pp. 1-45, `http://www.efsa.eu.int/science/ppr/ppr_opinions/1332/ppr_op_ej301_aquatic_ecotox_en1.pdf` (2005)

[2] United Kingdom Department of the Environment, Transport and the Regions and the United Kingdom Environment Agency, *Guidelines for Environmental Risk Assessment and Management*, `http://www.defra.gov.uk/environment/risk/eramguide/` (2000)

[3] United States of America Environment Protection Agency, *Risk Characterization Handbook*, US document number 100-B-00-002, `http://www.epa.gov/OSA/spc/pdfs/rchandbk.pdf` (2000)

[4] Elmegaard N and Jagers op Akerhuis GAJM, *Safety Factors in Pesticide Risk Assessment*, Danish Ministry of Environment and Energy National Environmental Research Institute, Tech. Report No 325 (2000)

[5] European Commission, *Technical Guidance Document on Risk Assessment in support of Commission Directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market, Part II, Chapter 3 Environmental Risk Assessment* (2003)

[6] Forbes VE and Calow P, *Extrapolation in ecological risk assessment: balancing pragmatism and precaution in chemical controls legislation*, BioScience **52**, pp. 249–257 (2002)

[7] Aldenberg T, Jaworska JS and Traas TP, *Normal species sensitivity distributions and probabilistic ecological risk assessment*, in [14] (2002)

[8] Solomon KR and Takacs P, *Probabilistic risk assessment using species sensitivity distributions*, in [14] (2002)

[9] United States of America Environment Protection Agency, *Glossary for the Integrated Risk Information System*, `http://www.epa.gov/iris/gloss8.htm` (2005)

[10] Congress of the United States of America, *The Food Quality Protection Act*, US Public Law 104-170, `http://www.epa.gov/oppfod01/fqpa/gpogate.pdf` (1996)

[11] Rosenbaum R, Pennington DW and Jolliet O, *An implemented approach for estimating uncertainties for toxicological impact characterisation.* In Pahl-Wostl et al (eds), Complexity and Integrated Resources Management, Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society (2004)

[12] Kodell RL and Gaylor DW, *Combining uncertainty factors in deriving human exposure levels of noncarcinogenic toxicants*, Annals of New York Academy of Sciences, **895**, pp. 188–195 (1999)

[13] Roex EWM, Van Gestel CAM, Van Wezel AP and Van Straalen NM, *Ratios between acute aquatic toxicity and effects on population growth rates in relation to toxicant mode of action*, Environmental Toxicology and Chemistry, **19**, 3, pp. 685–693 (2000)

[14] Posthuma L, Suter II GW, and Traas RP (eds), *Species Sensitivity Distributions in Ecotoxicology*, Lewis Publishers, Boca Raton, FL, USA (2002)

[15] European Food Safety Authority panel on Plant Health, Plant Protection Products and their Residues, *Mathematical and statistical basis of options for risk calculations*, appendix to [1], `http://www.efsa.eu.int/science/ppr/ppr_opinions/1332/ppr_op_ej301_aquatic_ecotox_appendix_en1.pdf` (2005)

[16] Seber GAF, *Linear Regression Analysis*, John Wiley (1977)

[17] O'Hagan A, *Kendall's Advanced Theory of Statistics Volume 2B: Bayesian Inference*, John Wiley (1994)